DOCUMENT RESUME

ED 262 092 TM 850 573

AUTHOR Worthen, Blaine R.
TITLE The Unvarnished Truth about Logic-in-Use versus Reconstructed Logic in Educational Inquiry. Draft.
PUB DATE 3 Apr 85
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142) -- Viewpoints (120)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Educational Researchers; *Heuristics; Inquiry; *Research Design; *Research Methodology; Research Needs; *Research Problems; Research Reports
IDENTIFIERS Reconstruction Technique

ABSTRACT

Kaplan (1964) describes the now classic distinction between "logic-in-use," the more or less logical cognitive style and procedures used by the educational scientist, and "reconstructed logic," the scientist's after-the-fact, explicit formulation and idealization of the logic and procedures. However, most technical documents read as smooth, polished, error-free examples of Kaplan's reconstructed idealizations, rather than reflecting the logic-in-use which would reveal many of the most important incidents in the drama of science, which reconstructed logic leaves safely behind the scenes. The author further contends that educational inquiry suffers from a collective inability to be candid in public reports. Three examples are offered to illustrate those factors perceived as influential in suppressing honest talk about trial and error efforts in research: (1) Scientific truth is often not the preferred way of "knowing" in education; (2) Sponsors of research, development, and education are often naive about science; and (3) Interorganizational and interpersonal politics often impede or prevent full and honest reporting. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Draft: Not for quotation or circulation
without author's permission.

ED262092

The Unvarnished Truth About Logic-in-Use versus Reconstructed Logic in Educational Inquiry

Blaine R. Worthen
Utah State University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Worthen, Blaine R.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 850 573

Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 3, 1985.

©B.R. Worthen. This material is not to be cited or reproduced without permission of the author.

The Unvarnished Truth About Logic-in-Use Versus Reconstructed Logic in Educational Inquiry

At first, I was complimented by the invitation to participate in this symposium, but that was before I noticed the Symposium sub-title, "Trials and Errors", and realized why I had been selected. During my twenty years as an educational researcher and evaluator, I have committed enough errors to fill a book, let alone a symposium paper. Indeed, I can think of very few colleagues more qualified to talk about "trials and errors" than am I. Of course, if I knew your work better, I might revise that opinion. But that is precisely my point -- I do not know your work well, any more than you really know mine. Further, I doubt that colleagues in any field of inquiry really know the work of other colleagues within that field, unless they happen to have worked together directly.

It is tempting to attribute our mutual unfamiliarity with one another's work simply to choices we make in what we read. I do know colleagues who have avowed that reading Webster's unabridged dictionary is more stimulating than reading most of the research and evaluation reports in our field. It is my contention, however, that we would not be much enlightened by exchanging printed reports of our research and evaluation studies, even if we could persuade one another to read them. Why? Because only part of what we do and think as researchers, developers, and evaluators ever is included in our printed reports. You and I would know only that portion of our colleagues' research orientation, behavior, and capability which they elected to reveal. Kaplan (1964) was speaking of precisely this fact when he made the now classic distinction between "logic-in-use", the more or less logical cognitive style and procedures used by the scientist, and "reconstructed logic", the scientist's after-the-fact, explicit formulation and idealization of the logic and procedures. As Kaplan puts it:

©B.R. Worthen. This material is not to be cited or reproduced without permission of the author.

> ". . . reconstructed logic is not meant to be merely a description of what is actually being done by scientists . . . but rather an idealization of scientific practice . . . The idealization may be carried so far that it is useful only for the further development of logic itself, and not for the understanding and evaluation of scientific practice. (pp. 10-11).

With this distinction in mind, think back to the last "scientific" or technical report you wrote. And the one before that. And still earlier. In how many of them have you unabashedly recorded the real "logic-in-use" which drove your work? Did you faithfully record in your reports all the false starts, the blunders, the compromises, and the wonderful hindsights which afflict all who try to follow the scientific tradition, however faithfully? If you did, you're in an overwhelming minority, for most technical documents read as smooth, polished, error-free examples of Kaplan's reconstructed idealizations, rather than reflecting the logic-in-use (or perhaps alogic-in-use, or even illogic-in-use) which would allow us to see many of the most important incidents in the drama of science, which reconstructed logic leaves safely (and unembarrassingly) behind the scenes.

Even those few souls who are compulsive about reporting their errors usually cloak their admissions in such technical language that only their "fellow-travelers" with equal methodological acuity would ever know a procedural sin had been committed. Conscience may be satisfied but truth no better served, as even the uncertain or conflicting findings that result from our own errors are often made to appear as if we had anticipated them all along. And yes, I am familiar with the "limitations of the study" section dutifully tucked into many research reports, especially masters and doctoral theses. But these highly stylized sections are usually pre-occupied with cautious statements about things like sample size, the dangers of extrapolating the results to groups with different characteristics, or the

relative advantages and disadvantages of using ANOVA on dichotomous data. Now all these are very useful concerns, yet "limitations" sections in our reports almost never highlight the really serious problems that plague all science but seldom appear in print.

Perhaps my portrayal is too cynical, but I contend that we do suffer in educational inquiry from a collective inability to be as candid in public reports of our work as we are over dinner, when among our few trusted confidents. But is this situation unique to the efforts of educationists? Hardly. Kaplan's criticisms were aimed at behavioral science, broadly. And Hans Selye, in his delightful treatise on medical research, begins his preface with these words:

> "The object of this book is to present certain problems of science through the portrait of one scientist -- the only one I know really well. It is not an attempt to picture the ideal scientist which the beginner should try to emulate . . . By speaking freely about all my views, I shall also reveal a great naiveness and ignorance on topics about which I should either know more or remain silent. These sections could easily have been suppressed to save my dignity. But there are many decorously retouched biographies and memoirs of much greater scientists; if you are to learn anything from my remarks, I feel they should be uncensored. . . But the point is not to describe what I should have been, but what I am." (Selye, 1964, ix).

Since reading Salye's portrayal of the real world of medical research, it has often struck me that the field of educational inquiry might profit if its inhabitants would speak as freely and frankly about their work and the personal difficulties they encounter in it. I am not calling for a communal catharsis or confessional, but rather for a realization that only through true candor and personal disclosure to one another can we ever hope to learn from one another's mistakes as well as from the rarer successes we are more eager to share. Albert Einstein made this argument to his colleagues in the physical sciences decades ago when he said, "The only resonable way to educate is to act as an example -- if one can't help it, [as] a deterring example." (Einstein, 1933, p. 40.)

These refreshing calls from luminaries such as Einstein, Kaplan, and Selye were compelling to me when I first began doing educational research and evaluation. I was younger then, certitude stood close to hand and everything seemed so much clearer. I could tell my students that we owed it to one another and to our field or inquiry to report boldly on our "logic-in-use", and to refrain from misleading and self-serving idealizations of our work in reports stripped of all error and uncertainty. But that was before I was punished, repeatedly, for publishing the whole story, rather than the half-true but better accepted and tidier rendition of research to which our educational community had become accustomed. Yes, that was before I learned that consumers of educational science, like consumers in several other facets of society, have only a very selective palate for the truth.

Having migrated back to the field of psychology the past few years, I have gained more insight into some of the interpersonal, interorganizational, and political factors which mitigate against honest disclosure of one's less exalted moments or less exemplary behaviors. I have also come to understand why, despite my penchant for publishing logic-in-use, I have gradually conformed more and more to the norms and only occasionally put in print in a research or evaluation report much "self-revelation" beyond the customary "limitations of the study" section. Gradually I have become more guarded, exposing my conceptual or methodological "Achilles Heels" only to the occasional trusted professional I am confident would understand or profit from it.

In the remainder of this paper, I would like to speculate about some of the factors which bode ill for researchers or evaluators who determine to publish their real errors and misjudgments openly. To illustrate these points, I would like to share several war stories -- tales true

"logic-in-use" which deserve to be heard before I die, or become a full-time administrator, which amounts to about the same thing.

What follows makes no pretense of conceptual certainty or even of logical categorization. Indeed, it is nothing more than "ruminations" that I am willing to share with professional colleagues in the hope it might be helpful. These accounts, with names, places, and organizations changed where necessary to protect my hide, are organized under three factors which I see as major disincentives to reporting of logic-in-use. Collectively these factors explain, at least in part, why research and evaluation reports are so ritualized and devoid of honest accounting of the imperfect, intuitive, and inconsistent groping that constitutes such a large portion of what happens under the rubric of inquiry. The examples given are intended to illustrate how I perceive those factors as influential in suppressing honest talk about trial and error in our collective enterprise.

1. Scientific Truth Is Often Not the Preferred Way of "Knowing" in Education

As with most of us, my first real foray into research was my master's thesis, conducted back somewhere just after Neanderthal times. It was a comprehensive study of the relative effectiveness of so-called discovery and expository methods of teaching math concepts, conducted with eight classrooms in each treatment. The study yielded significant differences that fueled the debate about task presentation variables that raged during the early 1960's. My advisor thought the study should be published, the editor of the Journal of Educational Psychology (JEP) agreed, and I was delighted to see my study cited by those whose conclusions happened to coincide with my data. The excitment I experierced in this first venture into research was largely responsible for leading me to pursue a doctorate in educational and psychological research.

But just as a little knowledge is a dangerous thing, so is too much. Two years into my doctoral training, a persistent statistics professor finally convinced me that the conservative use of classroom as the unit of analysis was a preferable choice to the more liberal use of student as the unit of analysis. This began to trouble me, for my master's study had used the student as the analytic unit, and I began to worry that: (1) I may have made a sophomoric error (which shouldn't have troubled me much, since I was still pretty sophomoric); and (2) someone may notice it.

Months passed, and I began to breathe easier. Maybe JEP readers didn't know much about units of analysis, I reasoned. But conscience prodded, as did Kaplan's troublesome notions. Finally, I decided the only conscionable choice was to re-analyze the original data, using classroom as the analytic unit. The re-analysis washed out most of the statistically significant differences in the earlier report. So I submitted a partial "retraction" to the JEP, in which I noted the earlier analysis was probably too liberal (since clearly some dependence exists within classrooms), while the re-analysis was probably too conservative (since clearly student performance within a classroom is unlikely ever to be completely dependent). In essence, I published the re-analysis, cautioned the readers as to the conclusions from both analyses, ended the article with a nimble "caveat emptor", and settled back with the uneasy realization that I could now be cited as debating myself in the literature.

The reactions were intriguing. Ellis Page took the time to write a very thoughtful letter applauding my willingness to re-examine my own work and publish results which corrected (or at least cautioned about) a possible error in my earlier study. Others were less kindly, expecially those whose own pedagogical preferences (and books on teaching methods) had been supported by

my earlier findings, but not by the re-analysis. One fellow, more miffed than most, sent a rather cryptic letter which read, "Thanks for the re-analysis; thanks for nothing. May a Bird of Paradise fly up your nose!" I didn't know much about Birds of Paradise, but somehow his wish struck me as uncongenial, and I wondered if my research career was off to an inauspicious beginning. But I did frame Ellis' letter and that helped feed my budding but bruised belief that truth is more important than accolades.

2. Sponsors of Research, Development, and Evaluation are Often Naive' About Science

My first conscious effort to portray the trials and errors of a research study in a written report was when I directed a study on research training for the American Educational Research Association (AERA). This was in the early 1970s, and we were attempting to adapt some of the military's task analysis procedures to our setting to learn more about the tasks and competencies needed in educational inquiry. This was terra incognito for me and the ten eminent researchers who served as a steering committee for the study. So we proposed that the U.S. Office of Education (USOE) fund AERA for a two-year study, with the first year to be a pilot study so we could develop and de-bug our task analysis methodology and map the terrain we wanted to examine more carefully the next year. The USOE liked the idea and funded the pilot year, with the second year to be contingent upon the success of the first.

We learned a lot that first year, most conspicuously how little we really knew about how to analyze what researchers do and what skills and knowledge they need to do it. In the process of learning, we tried many methods and techniques which didn't pan out, sometimes feeling like Edison who, when midway to his success, commented he now knew ten thousand things that didn't work. In retrospect, some of our trials and errors seemed a bit foolish, even to us, which only proved we were coming to see more clearly. We were also

having fun, working in virgin territory, and gradually becoming confident that our new research methods and techniques would work. By the end of the pilot year, we had a set of tools we knew would work to collect solid data during the key second year.

Then we submitted our report of the first year to USOE. I had suggested to the steering committee that we should describe all of our miscues, mistakes, and misbegotten ideas, along with our ultimate successes that led to the sound methodology we proposed for Year Two. My colleagues agreed heartily, feeling we owed it to anyone else trying to use similar task analysis procedures to warn of all the blind alleys, pitfalls, and dry holes we had encountered in our search, so they would not expend energy in the same vain pursuits.

The reaction of the staff of the USOE unit sponsoring our study was an interesting study in and of itself. Publicly, they were cool toward our report and non-committal about second-year funding. I subsequently learned they were privately dismayed at the collective incompetence of what they had been led to believe was a premier group of AERA's finest minds. Worse, they were shocked at our monumental stupidity in admitting our mistakes in print, thus tarnishing their enviable reputation of funding only competent researchers and flawless research. I learned of these private views at an AERA convention when I overheard them voiced by two people standing in front of me on the hotel's escalator. Recognizing two underlings in the USOE unit funding our study, I failed in my resolve not to eavesdrop, as they lamented that the likes of supposedly competent researchers like Jason Millman, Bob Gagne', Ellis Page, and Ernie Rothkopf could make so many errors in one piece of work. So much for their understanding of the scientific utility of "logic-in-use", I thought. So much also for our second-year funding, it suddenly occurred to me, unless we could educate some people quickly.

We tried. Our Task Force of eminent researchers explained patiently to the USOE staff about "logic-in-use", and when it became clear we were making no progress on that front, we switched to the less complex explanation of what a pilot study is intended to do. That worked no better. The irrefutable fact in their minds was that our report, with "sponsored by the USOE" emblazoned on the front, confessed to mistakes -- trials and errors, if you will -- that would be an embarrassment to their reputations, even if we didn't care about ours.

Finally, we had to appeal to the Branch Chief, and a resolution was reached. We would receive second year funding, plus a new project officer (who did turn out to be a great improvement). Not surprisingly, however, our pilot-year report was not conspicuously featured in USOE summaries of research. In fact, I never even saw one on a USOE waiting room coffee table. It simply disappeared.

Too bad, we thought, but went ahead with our second year, making few mistakes this time (thanks to Year One). In reporting our final results, our Task Force still opted to refer to the lessons learned in the pilot year, hoping to be of help to other researchers who may later follow our footsteps into the tangles of task analysis. Stubbornness is seldom rewarded, however, and the inclusion of even a more subdued and brief recounting of Year One trials and errors won us few friends at USOE. Worse, it may have helped guide our Year Two report into the same black hole that had swallowed up our pilot report. If any of you possess a copy of either report, hang on to it -- it's a real collector's item.

So, "strike two" for "logic-in-use". But I was still young and idealistic and not about to be dissuaded from my duty to science simply by a petty bureaucrat's distaste for the truth.

Enter "strike three". I accepted a position as Director of Research and Evaluation at a large non-profit contracting agency. One of my first tasks was to determine how to evaluate one of the large curriculum development programs underway in the agency at that time. After carefully examining all of the curriculum plans and interim products produced by this program, I realized that I did not have the faintest idea of what they were trying to produce or accomplish. Feeling conceptually bereft (after all, my doctoral minor in curriculum development was by then many years passed), I decided it might be wise to ask the developers just what they had in mind.

So I asked, but came away more confused than ever, for it seemed apparent that they knew little more about what they were trying to do than I did, and that was nil. Or, if they did know, they were not able to articulate it. The most tangible response I received was that the curriculum products were intended to teach creativity and problem solving. Nothing wrong with such general goals, I reasoned, although I had some trouble matching them with some of the products, such as the blank envelope that contained only a paper clip and small slip of paper on which was printed, "How many things can you think of to do with this paper clip?" Several suggestions came immediately to my mind, but in the absence of any criterion levels I was not certain whether my responses would be construed as correct.

Having had prior familiarity with several productive national curriculum projects (e.g., Biological Sciences Curriculum Study, PSSC, and SSEC), it occurred to me that I might be biased in favor of curriculum development efforts that followed a roadmap laid out by pre-specified objectives. So, wary of my own potential bias, I decided not to involve myself in the evaluation, leaving it to the team already assigned to conduct it. After all, who was I to say that the emperor had on no clothes.

When I later read the evaluation report that was submitted to the funding agency, The National Institute of Education (NIE), I realized that the evaluators had had as much difficulty finding the emperor's clothes as had I. For example, in reporting the findings of a field test of one of the curriculum products (which we shall call SPUR, as a pseudonym), they found that some users "liked" the product but that it made absolutely no difference on any of the agreed-upon outcome measures. Perhaps troubled by whether they were too imperceptive to discern the beauty of the emperor's new outfit, the evaluators went far beyond "logic-in-use" in becoming apologists for their own study. A quote from the summary of the final evaluation report should suffice to make the point.

> The study did not supply evidence from which either long-term or short-term performance changes can be inferred. The results do not necessarily indicate that SPUR was not effective in promoting change, but that a variety of shortcomings and limitations of the evaluation design did not allow for the identification and measurement of these changes. In the same sense, the lack of evidence as to the impact of SPUR on secondary audiences does not necessarily indicate that SPUR had no impact, but that this finding may be due to the particular set of variables that were chosen to be measured. (Disguised quote from actual report.)

What did the summary say? Nothing, really. Just that the lack of apparent effectiveness of SPUR might not mean that the emperor was really naked, but only that the evaluators lacked the visual acuity to discover his garb. Even with my professed proclivity for admitting our mistakes, I thought that these evaluators had gone too far in accepting responsibility for SPUR's failure to yield significant results.

Unfortunately, the funding agency didn't think they had gone far enough. The response of the NIE project monitor to whom the report was sent was blunt. The following quote is taken from the actual letter, with only the names changed to preserve anonymity.

> Dear John: Thank you for the four copies of the SPUR Impact Study. I hope that our future efforts will be structured so that statements such as those in the "summary" will not have to be made. Rather I hope that we will be able to say something positive about changes in important performances. I have heard so many good things about SPUR that I am disheartened to lack evidence to allow me to argue for its effectiveness.
>
> The issue here is straightforward. The best arguement that I can make internally at the Institute and externally with the Congress for funding agencies such as yours is that the products you produce lead to measureable changes for the good of American schools. Regardless of the positive "feelings" I get about SPUR, it appears we cannot justify all the effort in terms of performance criteria. That is a drawback. But one which I think we can overcome in future efforts. Sincerely, George. (Disguised quote from actual letter.)

George's point seems clear. Unlike his counterparts in the previous example, he would not have flinched at the sight of a report which contained mere "logic-in-use". On the contrary, he was not even satisfied with abject surrender, in which the evaluators in essence blamed themselves for the product's failure. I thought his suggestion about the future evaluation efforts was as subtle as a bludgeon -- if you want funding continued for your agency, then you had better collect data that conform to my conclusions!

We were unable to comply with his request, reminding him that we were evaluators, not fiction writers. But we resolved that our subsequent field tests would be the best possible. Better qualified evaluators were assigned to re-do the field test of SPUR and other products of this curriculum development effort, and these much better studies also failed to demonstrate that the curriculum yielded positive benefits. We reported the results honestly, along with a few trivial problems we had encountered, pointing out, however, that none of them materially changed the outcomes. Shortly thereafter, funding for the program was withdrawn, and the curriculum developers dispersed for other quarters. Unfortunately, so did most of the evaluators, for their funding had been withdrawn simultaneously. But then who ever promised evaluators career security. The best that one might hope for is

that "total truth telling" might lead to rewards in Heaven, if one is big on delayed gratification.

3. Interorganizational and Interpersonal Politics Often Impede or Prevent Full and Honest Reporting

Up to this point, I have commented on two disincentives for revealing one's trials and errors in printed reports, both of which would suggest that it is necessary and noble to insist on total honesty on every occasion. "Tell it like it is" is a grand ideal, but in some contexts the political realities are such that complete revelation would be the offspring of sheer stupidity. Profoundly complicated interpersonal and interorganizational political forces can combine to form such an explosive context that even a simple reporting of "logic-in-use" would have devastating side effects. Perhaps a few simple illustrations might serve to make the point.

One such example is the evaluation of a controversial elementary school reading program some colleagues and I conducted. The controversy resulted from split scheduling, where half of the first graders came to school at the normal time but departed an hour early, while the second half came an hour late and remained until the normal close of school. As a result, first-grade teachers were able to teach reading in the first and last periods of the day, with only half of the class present at these hours, thus providing greater individualized instruction.

Surprisingly, the program was vigorously opposed by a small, but well-organized, group of parents. The school teachers, principals, and central office administrators were enthusiastic about the program and rallied to its defense. The parents countered with accusations that the staggered scheduled was unsafe for their younsters (since they were unable to walk to and from school with older siblings), that teachers were being paid a full day of instruction but were teaching only 50% of normal load for two hours of each

day, and that children were being shortchanged in their education, since parents could hardly be expected to provide adequate substitutes at home for the loss of one hour of schooling each day.

The district had already conducted one survey of parents, but this was discredited by the opposition group as biased, because it was conducted by district personnel, and invalid because of some minor technical problems which did not substantially alter any of the outcomes. Finally, the Board of Education called us to conduct an external, unbiased opinion survey of parents in the community to determine whether there were any wide-spread parental concerns about the program.

We conducted the study, careful to include input from the district and the opposition group, through a balanced steering committee which approved all survey questions, sample selection, and so on, to ensure fairness to all points of view. The results showed that parents were overwhelmingly favorable toward the program, on virtually every dimension. Unfortunately, on two of the seventeen questions posed in the survey, it appeared that one surveyor may have encoded the data incorrectly. It was one of those instances when you have the uneasy feeling that an error may have led to distortion of the data, but you cannot be certain because adequate safeguards had not been in place to double check and cross validate all data collection. Normally, I would be compulsive about reporting such a possible problem, even though it would not have substantially altered the outcome on even those two questions. In the present instance, however, I opted to make no mention of this possible problem in the report.

Why? Simply because it had also become clear to us during the study that the real objection of the opposition group to the split-schedule program was not for any of the reasons they publicly pronounced. Indeed, after one

difficult steering committee meeting, there was a heated exchange between one of the leading opponents to the program and a teacher who was stoutly defending it. In the midst of the shouting, it became apparent that the opposition's real objection to the staggered schedule was the fact that an extra hour of parental supervision was required for each child, thus seriously cutting into the bridge, tennis, and other social activities of those who had organized the opposition to the program. Furthermore, this spokesman for the opposition expressed hope that the forthcoming survey would contain enough flaws that they could cast doubt on the findings, as they had on the district's survey.

Now I have nothing against bridge or tennis, and I can even argue that babytending can be construed as an appropriate function for schools, but I am opposed to aiding and abetting anyone so unprincipled as to launch spurious accusations at a program, expend taxpayers money to study it, and then lie in wait to discredit the study, all in the cause of personal convenience. So, no mention of our uneasiness about the two questions appeared in our report. Indeed, the report contained only the pure distillation of "reconstructed logic", along with beautiful graphics, which revealed that a great majority of parents favored the staggered schedule, except in one elementary school which was, not coincidentally, located closest to the Country Club tennis courts. The opposition found no flaws, for we were determined that the methodological pimples in our study would not be portrayed as terminal acne, and the school Board approved continuation of the program.

Were we correct in taking that course? Frankly, I don't know, for I know that we allowed our values to enter into the decision of how we handled the report; but I would do it again.

A similar circumstance existed in a recent national evaluation we were asked to conduct for the Humane Society for the U.S. They had developed and piloted a set of humane education curriculum materials and asked us to determine whether the curriculum was effective in improving children's humane treatment of animals, attitudes toward animals, kindness toward other children, and the like. Two school districts in a western state expressed interest in using the materials and agreed to participate in a district-wide quasi-experimental field test to enable us to pilot our instruments and data collection methods. Everything was going well until we walked into a scheduled meeting with a two-district steering committee to arrange final details for the study. To our surprise, we found ourselves confronted by a room full of angry local farmers, who had been informed by the County Farm Bureau and State Farm Bureau officials that the districts were about to adopt a curriculum which would threaten American agriculture. We quickly learned that the U.S. Farm Bureau and the Humane Society of the U.S. had been engaged at the national level in a two-year running battle concerning these materials. The U.S. Farm Bureau had warned each state and local affiliate to beware of allowing the intrusion of these materials into their schools. They claimed that the materials taught vegetarianism, condemned factory farming as cruel, and advocated pet population control concepts such as "spaying" or euthanasia which were certain to lead to rising generations of abortionists or mercy killers. They also argued that exposure of children to the materials would lead to eventual reduction of factory farming, thus threatening the very basis of American agriculture and our economy by driving food prices up to heights beyond reach of the average citizen.

We did not doubt the sincerity of those good honest farm folk who had gathered to discuss their concerns with us -- after all, there is a certain

sincerity about poised pitchforks. It was not the local farmers who really concerned us, however, but rather the guile of those who had orchestrated their being there. Representatives of the State Farm Bureau were standing by both doors passing out selected pages copied from the curriculum guides, with words like "communist plot" written in the margins, with arrows drawn to circled passages. I found it amusing when I realized one circled reference was to a statement in the materials describing that bee colonies acted as "collective societies", but it was clear that few in the room shared my amusement.

We tried to discuss the issue rationally, pointing out that our position as evaluators was neutral and that, if the materials were as bad as the Farm Bureau believed, they should be delighted to see them evaluated and duly exposed as a thinly veiled plot to take over American society. We also pointed out that our analysis of the materials had neither revealed any support for vegetarianism nor condemnation of any particular farming practices. In fact, to our innocent eyes, the materials were seemingly aimed only at sensitizing youngsters to the view that no unnecessary pain or suffering should be inflicted on infra-human species. We invited the Farm Bureau folks to show us what we had overlooked, but they declined.

To shorten the story, reason did not prevail. Political demagoguery did. Local school officials were threatened with loss of bond elections and possible loss of employment. Although beleaguered, the school administrators were willing to stick by their original committment to conduct the study, but we thought it inadvisable to continue in the face of such political opposition. Common decency suggests that it is unkind to allow your friends to be spitted on five-tined hayforks, if there is any reasonable alternative. In this case there was. We simply regrouped, made contacts with school

districts in Connecticut and California where adequate cross sections of rural and urban schools could be found, but where the state and local farm bureaus seemed less fond of McCarthyism, and we proceeded with the study.

The study went well, with relatively few minor hitches and glitches, and the results were provocative, showing that the curriculum had substantial impact on children in kindergarten and first-grade, with significantly less impact on children thereafter. Indeed, the study proved of interest to child development specialists, as well as to those with interest in humane education. But we had learned something besides raw fear from our encounter with the farm folks, namely that our research report would be heavily critiqued by the best "hired guns" the U.S. Farm Bureau could afford. So, we consciously chose to omit mention of any of our procedural faux pas, excepting two which were central, since they invalidated one of our instruments in the Connecticut sites, causing a drop in our overall sample size which rendered the results on that instrument somewhat equivocal. Should we have included mention of the other minor methodological miscues in our study? Again, I cannot respond with certainty, but omitting them from our report seemed the right thing to do at the time. After all, who wants to present a standing target for a hired gun, especially one who is employed by someone who has already branded you as a pinko.

If time permitted, I would provide other examples, such as the evaluation of an interesting state-wide program in a seaboard state, in which I first learned that the role of evaluator is not free of physical perils. We also learned that longshoremen in that state have a keen social consciousness. Having agreed to support the stance of a state employees union on a particular educational program, the longshoremen were eager to prove that they were something more than a bunch of pug-uglies from the dock -- and they would punch you out to prove it. We were careful with that evaluation report.

There was also the time when I was called on to evaluate an evaluation syposium conducted by eminent colleagues who were kind enough to loan me their secretary to assist with my on-site typing, data tabulation, etc. It took me until the third day of the syposium to realize that I had been had. My loan-lease secretary was loyally tipping her bosses in advance to every kind of observation our team had planned, questions we would be asking in interviews with participants, and so on. I might have reported on the "logic-in-use" in that study, if I had been certain just whose logic had actually been used to guide its direction. Instead, I simply jettisoned the "tainted" data and reported that which was uncontaminated, which I judged still enough to provide adequate conclusions about the symposium. Sure, I could have skewered my colleagues for shady behavior, but to what end? They were already mortally embarrassed and repentant. Why kick someone who is already down. Besides, they were evaluators, so I sensed being evaluated was a new and unsettling experience for them.

4. Conclusion

Let me try to pull this rambling presentation back into focus with a few concluding comments.

First, I am not championing sloppiness in research or evaluation. Suggesting that we talk openly about our errors is not tantamount to suggesting that we be casual in making them. Especially, there is little excuse for repeating errors made previously by ourselves or others, which is, after all, the prime reason for reporting our trials and errors in public presentations of our work.

Second, I believe that we have an obligation to try to educate educators and colleagues in related fields as to what science really is. Somehow, we have to help people realize that the notion of unflawed metholology is nothing

more than unfounded mythology. Somehow, we need to stop the charade of precision and perfection in studies of educational phenomena which are so complex that our best efforts at understanding them are fumbling. Only by following the counsel of Kaplan, Seyle, and others who understand that science is only as infallible as the scientist can we hope to truly come to know one another's work in ways that will mutually help us to advance our knowledge.

Finally, I obviously do not believe that it is invariably appropriate to reveal every mote in our methodological mistakes. Common sense -- not to mention common decency -- sometimes dictates that conscience, not compulsiveness, determines how open we dare be in discussing our errors wihout doing harm to those we hope to help.

Yet, in the main, I believe we do ourselves and our collective field of endeavor a disservice by our persistent and pervasive tendency to stylize and reconstruct our research, development and evaluation efforts so as to purge from them all trace of doubt and error. Let's quite kidding ourselves. Let's learn to admit our mistakes and help educate those currently too innocent to recognize that all of our work is flawed.